Quantum software for linear photonic simulations

B. Opanchuk[1], L. Rosales-Zárate[1], M. D. Reid[1], P. D. Drummond[1,2]
[1] *Centre for Quantum and Optical Science, Swinburne University of Technology, Melbourne 3122, Australia and*
[2] *Kavli Institute for Theoretical Physics, UC Santa Barbara, USA*
(Dated:)

The search for new, application-specific quantum computers designed to outperform any classical computer is difficult and ambitious, yet strongly driven by the ending of Moore's law and the great quantum advantages obtainable. Boson sampling photonic networks are highly promising examples of this, with preliminary experimental demonstrations and strong potential for obtaining the first quantum computer to solve problems believed classically impossible. We develop novel complex phase-space software for simulating these photonic networks, and apply this to two outstanding challenges: boson sampling experiments, and calculating the decoherence of quantum Fourier metrology. Our stochastic techniques give sampling errors orders of magnitude lower than experimental measurements of correlations, for the same number of samples. We show these techniques remove systematic errors in previous algorithms for estimating correlations, with order of magnitude improvements in errors in some cases. We obtain a scalable channel-combination verification strategy for boson sampling, and simulate a hundred qubit Fourier transform interferometer — far larger than possible in conventional calculations — demonstrating that this quantum metrology technique is robust against decoherence.

INTRODUCTION

Linear photonic networks are used to implement novel quantum protocols, including boson sampling [1, 2] and future, high accuracy, quantum Fourier interferometers [3]. In linear photonic experiments one prepares an arbitrary M-mode bosonic state $\hat{\rho}$, which is input into a passive linear optical multimode device, followed by a measurement on the output [4]. For single photon inputs into multiple channels, the generation of the output photon-count bit-stream is conjectured [2, 5] to be an exponentially hard problem [6], and therefore not feasible on a classical computer for large M.

These remarkable experiments [3, 4, 7–12] have the goal of demonstrating computations which are thought to be impossible on classical computers, and lay the foundations for new quantum technologies. Nevertheless, one must have tools to analyse them. Here we derive a novel hybrid computational and analytic approach to allow the analysis and verification of such networks. For both theory and experiment, we emphasise the crucial role of sampling errors in determining complexity and computability.

Linear photonic devices are defined by an $M \times M$ mode transformation matrix, which can include losses. Boson sampling experiments [7–13] have an initial N-photon state $|\boldsymbol{n}\rangle$, where $\boldsymbol{n}$ is a vector such that $N = \sum n_j$, $n_j = 0, 1$, and $|n\rangle$ is the number state basis. The output photon numbers, $\boldsymbol{n}'$, are the observables. The permanent-squared of the sub-matrices defined by the input and output modes [3, 14, 15] gives the probability of measuring one photon in each preselected output mode [3]. The optimal classical techniques for calculating this probability scale as $N2^N$ [16] operations, and become rapidly infeasible above $N = 50$ [17]. To understand the reason for this, we note that the permanent of a square matrix $\boldsymbol{U}$ is defined as a sum over all permutations σ of the set of indices $\{1, \ldots, N\}$, $\mathrm{perm}(\boldsymbol{U}) = \sum_\sigma \prod_i U_{i\sigma(i)}$. The number of such permutations rises exponentially with N. This complexity occurs naturally in a photonic network, driven by the enormous number of possible interference paths that a photon can take.

It is an open problem how to verify that boson sampling experiments work correctly, due to the extraordinary hardness of matrix permanent calculations. We propose a new strategy for verification at large network size, together with simulation methods that can treat arbitrary quantum inputs and outputs, as occur in practical applications. These novel techniques do not replace, but rather complement quantum hardware. Our methods do not solve hard problems in polynomial time, which is conjectured to be impossible on classical computers, but rather they provide avenues for verifying that the hardware-generated solution is valid.

We simulate linear photonic experiments using "quantum software", by transforming this problem into an exact expansion on non-classical phase space [18]. The importance of using a nonclassical phase-space is that it allows unbiased estimates of the experimentally measured probabilities, even at very large sizes. For measurements of photonic correlations, we show that this method scales *better* than an experiment, in terms of sampling error. In addition, it removes systematic errors that can occur when estimating correlations using previously known algorithms for calculating approximate permanents.

Utilising these techniques, we propose an analytic signature for validating claims to solve Boson Sampling, using a hierarchy of combinations of N-th order correlations. Such validation is an essential part of any solution to exponential complexity, and involves sums over exponentially large numbers of sub-permanents present in the experiments, in order to obtain significant count-

rates. The analytic theory itself is verified numerically, which we have carried out efficiently in a single run of our simulations. The results far exceed previous limits [17] on such calculations. Our approach can represent any input state or output measurement. This is essential for understanding how this technology can be scaled to large sizes with imperfect sources [19], or used in different applications [20].

RESULTS

Complex P-representation

We first explain our "quantum software" method, which uses the complex P-distribution. This is a quantum phase space expansion over a basis of coherent state projectors [18]. We use this method to "simulate" linear photonic experiments, evaluating the measurable photonic moments by probabilistically sampling over specifically selected contours in the higher dimensional complex space $\mathbb{C}^M \times \mathbb{C}^M$. In the complex P-representation [18], the input quantum density matrix is represented by an integral over a closed contour C enclosing the origin in a multi-dimensional complex plane:

$$\hat{\rho}^{(\mathrm{in})} = \oiint_C P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) d\boldsymbol{\alpha} d\boldsymbol{\beta}\,.$$

Here, the quantum operator basis $\hat{\Lambda}$ is a set of coherent state projectors

$$\hat{\Lambda}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) = \frac{\|\boldsymbol{\alpha}\rangle\left\langle\boldsymbol{\beta}^*\right\|}{\left\langle\boldsymbol{\beta}^*\right\| \boldsymbol{\alpha}\rangle}\,,$$

defined in terms of un-normalised Bargmann-Glauber coherent states $\|\boldsymbol{\alpha}\rangle$, where

$$\|\boldsymbol{\alpha}\rangle \equiv \prod_k \left[\sum_{n_k} \frac{\alpha_k^{n_k}}{\sqrt{n_k!}}\left|n_k\right\rangle\right]\,.$$

For the purposes of the sampling, it is only necessary to know that α_k and β_k are complex numbers, and that the phase space method will generate the correct quantum moments. We choose to integrate around a circular contour with a complex-valued weight $P\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)$. The over-completeness of the coherent states in quantum mechanics means that there is more than one way to choose the contour C, and in particular the circular radius can be varied. It is an essential feature of our method that this flexibility enables us to tailor the representation to optimise the sampling for different tasks.

To perform simulations using this approach it is necessary to have a representation, $P\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)$, of the initial state. This distribution exists for an arbitrary initial density matrix, either pure or mixed, with any coherence properties and photon numbers [18]. In this paper, we focus on the standard case: a pure state in which the k-th channel has an n_k-boson input. The usual boson sampling experiments send one photon into each of a set S of input modes, so that $n_k = 0, 1$. There is more than one way to represent these states, which are illustrated below.

Any bosonic input density matrix $\hat{\rho}^{(\mathrm{in})}$ is changed by transmission through a linear network to an output density matrix $\hat{\rho}^{(\mathrm{out})}$. To calculate the effect of this transmission, we introduce a transmission matrix $T = \sqrt{t}U$. This combines the unitary mode transformation U of the network, with an absorptive transmission coefficient t. One can equivalently consider a larger unitary that includes loss channels, such that T is the sub-matrix for the accessible channels that are measured.

The effect of the transmission matrix T on the phase-space distribution is straightforward, owing to the normal ordering property of the representation. It simply transforms the coherent amplitudes in a deterministic way [21], such that $\boldsymbol{\alpha}^{(\mathrm{out})}$, $\boldsymbol{\beta}^{(\mathrm{out})} = T\boldsymbol{\alpha}$, $T^*\boldsymbol{\beta}$. The resulting output density matrix is therefore a contour integral with the same weight, but a modified projector:

$$\hat{\rho}^{(\mathrm{out})} = \oiint_C P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}\left(T\boldsymbol{\alpha},\, T^*\boldsymbol{\beta}\right) d\boldsymbol{\alpha} d\boldsymbol{\beta}\,.$$

Any output number correlation is given by computing moments of the input P-function, using the output number variable $n_k^{(\mathrm{out})}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \alpha_k^{(\mathrm{out})}\beta_k^{(\mathrm{out})} = \left(\sum_j T_{kj}\alpha_j\right)\left(\sum_j T_{kj}^*\beta_j\right)$. A typical observable in these experiments is an arbitrary normally ordered quantum correlation of k-th mode operators, $\hat{n}_k \equiv \hat{a}_k^\dagger \hat{a}_k$, belonging to a set of output modes S'. This is given by the following expression,

$$\begin{aligned}\left\langle \prod_{k\in S'} \hat{n}_k \right\rangle_Q &= \oiint_C P(\boldsymbol{\alpha},\boldsymbol{\beta}) \prod_{k\in S'} n_k^{(\mathrm{out})}(\boldsymbol{\alpha},\boldsymbol{\beta}) d\boldsymbol{\alpha} d\boldsymbol{\beta} \\ &= \left\langle \prod_{k\in S'} n_k^{(\mathrm{out})}(\boldsymbol{\alpha},\boldsymbol{\beta}) \right\rangle_P\,.\end{aligned}$$

As the distribution is not unique (the choice of C being flexible), one can choose different representations of the input state. These lead to different strategies for using random sampling methods, as well as different sampling weights. In this work we use two kinds of representations tailored for different scenarios, described below: a continuous sampling method (VCP) and a discrete method (QCP). Which is preferred depends on the results being calculated.

Continuous complex P-representation: VCP results

This section describes a general continuous complex P-type distribution, while the next section describes a

discrete representation which restricts the initial state and measured correlations. We start from a continuous analytic complex P-distribution for the initial state

$$P\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)=\prod_{k}\left[\left(\frac{n_{k}!}{2\pi i}\right)^{2}\frac{e^{\alpha_{k}\beta_{k}}}{\left(\alpha_{k}\beta_{k}\right)^{n_{k}+1}}\right], \tag{1}$$

where n_k is the photon count in the mode k. For $n_k = 0$, the single pole at the origin means that one can replace the input variable by its vacuum value of $\alpha_k = \beta_k = 0$ in any average, so that $P\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)=\prod_{k}\delta\left(\alpha_{k}\right)\delta\left(\beta_{k}\right)$ for the vacuum modes.

A complex contour is used for nonzero boson number inputs. Rather than integrate this analytically, it is numerically sampled for the complex-P representation defined in Eq. (1). It is convenient to integrate on a circular contour C of radius r, using an angular measure, so that $\alpha_k = rz_k = r\exp\left(i\phi_k^{(\alpha)}\right)$ and $\beta_k = r\tilde{z}_k^* = r\exp\left(-i\phi_k^{(\beta)}\right)$ respectively, where the coherent modulus r is chosen to minimise the sampling error.

The phase variables can be understood intuitively on defining $\phi_k = \left(\phi_k^{(\alpha)} + \phi_k^{(\beta)}\right)/2$ and $\theta_k = \phi_k^{(\alpha)} - \phi_k^{(\beta)}$, where ϕ_k is the classical phase, and θ_k is a nonclassical phase which only exists when the quantum state has nonclassical features. In this case we get:

$$\begin{aligned}\hat{\rho}=&\prod_{k}\left[\frac{\left(n_{k}!\right)^{2}}{4\pi^{2}r^{2n_{k}}}\int_{-\pi}^{\pi}d\theta_{k}\int_{-\pi}^{\pi}d\phi_{k}\right.\\&\left.\times\exp\left(r^{2}\cos\theta_{k}+i\left(r^{2}\sin\theta_{k}-n\theta_{k}\right)\right)\hat{\Lambda}\left(\alpha_{k},\beta_{k}\right)\right].\end{aligned} \tag{2}$$

It is clearly possible to separate the real and imaginary parts of the exponential. We use random probabilistic sampling for the real part. The resulting distribution is called a circular von Mises complex-P distribution (VCP), since the weight around the contour has a von Mises probability distribution [22]. Once the random phase angle is randomly chosen, the imaginary part is included as an additional complex weight. The full sampling algorithm is described in the Materials and Methods section.

In general cases it is most efficient to use this method, with a finite radius and von Mises sampling, especially for low-order correlations. This approach also provides a link with contour integral methods for matrix permanents [23].

Discrete complex P-representation: QCP results

Suppose that the initial photon number is bounded, for example with a fixed input boson number. We now introduce a discrete sampling method, which we term the discrete or qudit complex P-representations (QCP) [24]. This construction is useful in the limit of $r \to 0$, which is expanded using d coherent phases distributed on a circle. This is equivalent to a d-dimensional qudit, with initial occupation numbers in each mode of $n = 0, \ldots d-1$. This approach unifies quantum optical representation theory [18] with discrete sampling permanent approximation methods [25].

In this approach we consider a discrete set Q of coherent amplitudes defined as

$$\begin{aligned}\alpha^{(q_j)}&=rz^{q_j}=r\exp\left(iq_j\phi\right),\quad q_j=0,\ldots d-1\\ \beta^{(\tilde{q}_j)}&=r\left(\tilde{z}^{\tilde{q}_j}\right)^*=r\exp\left(-i\tilde{q}_j\phi\right),\quad \tilde{q}_j=0,\ldots d-1\end{aligned}$$

The density matrix is expanded in coherent states with a discrete summation:

$$\hat{\rho}=\frac{1}{d^{2M}}\sum_{\mathbf{q},\tilde{\mathbf{q}}}P(\mathbf{q},\tilde{\mathbf{q}})\hat{\Lambda}(\mathbf{q},\tilde{\mathbf{q}}),$$

where $\hat{\Lambda}(\mathbf{q},\tilde{\mathbf{q}})=\left\|\boldsymbol{\alpha}^{(\mathbf{q})}\right\rangle\left\langle\boldsymbol{\beta}^{(\tilde{\mathbf{q}})*}\right\|$, noting that these coherent states have unit norm in the limit of $r \to 0$. Using this expansion, a complex qudit P-function P_Q always exists for a given input density matrix $\widehat{\rho}$, where:

$$P_{Q}(\mathbf{q},\tilde{\mathbf{q}})=\sum_{\boldsymbol{n},\boldsymbol{m}}\left\langle\boldsymbol{m}\right|\widehat{\rho}\left|\boldsymbol{n}\right\rangle\prod_{j}r^{-(n_{j}+m_{j})}e^{i\phi(n_{j}\tilde{q}_{j}-m_{j}q_{j})}\sqrt{n_{j}!m_{j}!}.$$

In the simplest binary, or qubit, case where $n_k = 0, 1$ (so $d = 2$), we consider $\phi_k = \{0, \pi\}$, which implies that $\alpha_k = \pm r$. In the qudit case, for a single-mode Hilbert space dimension d, we consider d random discrete phases. Here ϕ_k can have values of $(0, \ldots d-1)\times\phi$, with $\phi = 2\pi/d$, allowing occupation numbers for mode k up to $n_k = 0, \ldots d-1$. We note that such discrete sampling reduces to continuous sampling in the limit of $d \to \infty$.

This method is more efficient for highest N-th order correlations, and especially useful in the case of boson sampling interferometry. Details are given later in this section, as well in the Materials and Methods section.

Error scaling

There are close similarities between experimental measurements and the use of "quantum software". In both cases there is a sampling error caused by the fact that one must calculate results for correlations or moments using a finite number of samples. The complexity is proportional to the number of samples. It is therefore crucial to know how the average sampling error scales with the number of active channels N, since this determines the computational time required.

To make useful predictions about experiments, perfect accuracy is not essential. It is only necessary to calculate permanents with better than experimental errors. These in turn depend on the average scaling behaviour of N-channel coincidences, which can be calculated analytically. In a boson sampling experiment, the probability

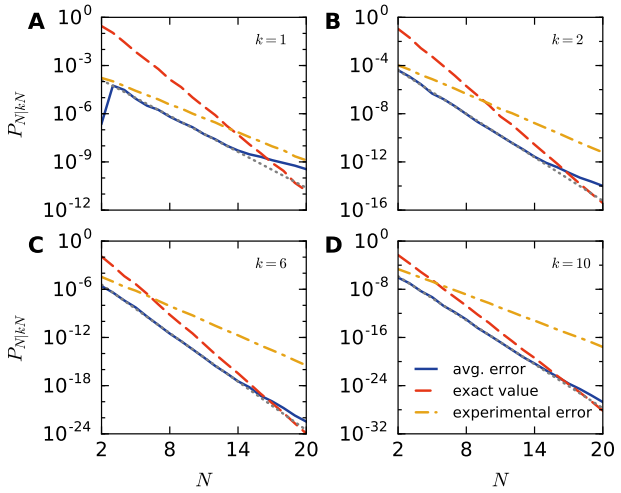

FIG. 1. **Scaling of errors for the coincidence rate**. **(A-D)** Estimation of the error for coincidence rate $P_{N|kN}$ as a function of N using the QCP method with a random discrete phase $d \to \infty$. For each point we have used $N_m = 100$ random unitary matrices and $N_e = 100$ ensembles of $N_s = 10^5$ samples each. Dotted red line corresponds to the average for the exact value of $P_{N|kN}$. Solid blue line is the average error compared to the exact value E. Dash-dotted yellow line denotes the estimated experimental error $\sqrt{\langle P\rangle_{e,m}/(N_s N_e)}$. Dotted grey line is a linear fit of the error. Here we consider: (**A**) $k = 1$, (**B**) $k = 2$, (**C**) $k = 6$ and (**D**) $k = 10$.

of finding N photons in the M output modes of a linear optical network is given by the permanent of an $N \times N$ sub-matrix of an $M \times M$ matrix, $\left(\text{perm}\left(U_{N|M}\right)\right)$ [9]. We denote the average over all unitary transformations of this permanent as $\left\langle \left|\text{perm}\left(U_{N|M}\right)\right|^2 \right\rangle_U \equiv P_{N|M}$. This has a known scaling law given by [26, 27]:

$$\log P_{N|M} = N\epsilon\left(k\right) + O(\log N), \tag{3}$$

where $\epsilon\left(k\right) = k\log k - (1+k)\log(1+k)$ and $k = M/N$ is the channel ratio.

In any experiment with T trials, the average Poissonian measurement error due to shot-noise is $\sigma = \sqrt{P/T}$, which therefore scales as $\log\sigma = N\epsilon\left(k\right)/2 - (\log T)/2 + O(\log N)$. To verify an experimental probability, one must have a theoretical error less than this experimental sampling error.

Figures 1A-D shows the exact modulus squared of the permanent of an $N\times N$ sub-matrix, denoted by $P_{N|kN}$, as well as the average deviation of the sampled result from the exact value E, defined as $E = \langle|P - \langle\tilde{P}\rangle_e|\rangle_m$, where $\langle\tilde{P}\rangle_e$ is the quantum software ensemble average. All results are averaged over a finite unitary ensemble of matrices $\langle\rangle_m$. We have also plotted the Poissonian experimental error, which is asymptotically *larger* than the error of our simulations for an identical sample number. This shows that the computational error using our method has a *better* scaling than experiment, for the same number of measurements or samples.

This is a necessary requirement for any verification method, since without this favourable scaling, the calculation of expected correlations would take exponentially longer than the experiment itself. Our computational sampling errors reduce rapidly with N — much faster than experimental error reductions — but the average permanent values reduce even faster. This makes direct permanent measurement problematic at large N. In other words, at sufficiently large N the count rate for any individual set of channels goes rapidly to zero, so that the experimental measurement of any individual correlation or moment is impractical in a finite time. Before explaining how this experimental problem is solved, we explain in greater detail how these advantageous theoretical scaling properties are obtained, and how our results compare to other methods that could be utilised.

Sampled calculations

In both the continuous phase and discrete phase approach, we use independent random samples of $\mathbf{z}, \tilde{\mathbf{z}}$, which correspond to scaled versions of $\boldsymbol{\alpha}, \boldsymbol{\beta}$ in the original complex P-representation. It is essential, in order to obtain an unbiased estimate of the absolute value squared, that one uses two sets of size $\tilde{N}_s = N_s/2$ of *independent* conjugate variables, $\mathbf{z}, \tilde{\mathbf{z}}$. Each approximation of the permanent, which we denote as $p(\mathbf{M}, \mathbf{z})$, is a function of the sub-matrix $\mathbf{M}$ and a noise vector $\mathbf{z}$.

Both methods give an unbiased estimate of the modulus-squared of the permanent, and are applicable to any type of matrix permanent calculation, although the error scaling with matrix size depends on the algorithm and type of matrix. We use $\langle\rangle_{N_s}$ to denote a stochastic expectation value over N_s random samples of a stochastic vector, so that:

$$\langle p(\mathbf{M},\mathbf{z})\rangle_{\tilde{N}_s} \equiv \frac{1}{\tilde{N}_s}\sum_{j=1}^{\tilde{N}_s} p(\mathbf{M},\mathbf{z}^{(j)})\,.$$

The permanent squared is initially approximated as a function of a set of samples, $\mathbf{Z}, \tilde{\mathbf{Z}}$, each composed of $\tilde{N}_s$ random vectors:

$$\tilde{P}(\mathbf{M},\mathbf{Z},\tilde{\mathbf{Z}}) \equiv \Re\,\langle p(\mathbf{M},\mathbf{z})\rangle_{\tilde{N}_s}\,\langle p(\mathbf{M},\tilde{\mathbf{z}})\rangle^*_{\tilde{N}_s} \tag{4}$$

The noises variables factors into two terms which are independent but conjugate on average. The factored terms give independent estimates of the permanent and its conjugate, so their product is an unbiased estimate of the modulus squared. We take the real part to impose the constraint that the final result must be real.

Sampling error estimation

In any calculation that uses random sampling, it is essential to have a statistical procedure for estimating the errors. This gives a theoretical error-bar, which indicates how accurate the calculation is. In all our stochastic calculations, we use sub-ensembles, each including N_s random noise vectors. Since the sub-ensembles have a large number of independent noise terms, their averages have a Gaussian distribution for large N_s, by the central limit theorem. This gives us independent estimates of the permanent and its conjugate.

Combining these in pairs gives us an unbiased estimate of the permanent squared. We then take an average over N_e such estimates to give an overall average that is used for estimation, together with error estimates:

$$|\mathrm{perm}(\mathbf{M})|^2 = \left\langle \tilde{P}\left(\mathbf{M}\right)\right\rangle_{N_e} \equiv \frac{1}{N_e}\sum_{i=1}^{N_e} \tilde{P}(\mathbf{M}, \mathbf{Z}^{(i)}, \tilde{\mathbf{Z}}^{(i)}).$$

This allows us to estimate sampling errors using statistical methods valid for Gaussian distributions, giving accurate estimates for the variance in the mean, as plotted in the figures.

In many of these results, we are interested in typical properties of random unitaries. In these calculations, we therefore obtain a set of randomly chosen unitary matrices according to the Haar measure over the unitaries, to give a *third* level of random averaging. We repeat the above process for each unitary matrix, to get unitary averages over the sampling errors that occur in simulating typical random unitary matrices.

There is a close relationship between the QCP method with qubits for the permanent squared, and the older Gurvits method for permanent approximation [25], in the case of an N-photon input and output. Both methods employ random sampling over binary numbers. We extend this to qudit and continuous cases, which has advantages for the combined calculations given below. The earlier method provides an unbiased estimate of the permanent. However, it introduces systematic errors when computing the permanent squared, owing to effects of the finite distribution width, and very large sampling errors when estimating combinations of permanents.

Fig. 2A-B shows the systematic errors of the modulus-squared of the permanent using the older method and the novel QCP representation with $d \to \infty$ for a ratio $k = 4$, and 10×10 and 20×20 unitary matrices respectively. These figures show that the Gurvits algorithm has a statistical bias for the permanent squared which is of an order of magnitude larger than the standard deviation. This statistical bias is reduced as the number of samples is increased, but the samples required rapidly become impractical as the matrix size N is increased. A description of the sampling for these plots is given on the Material and Methods section.

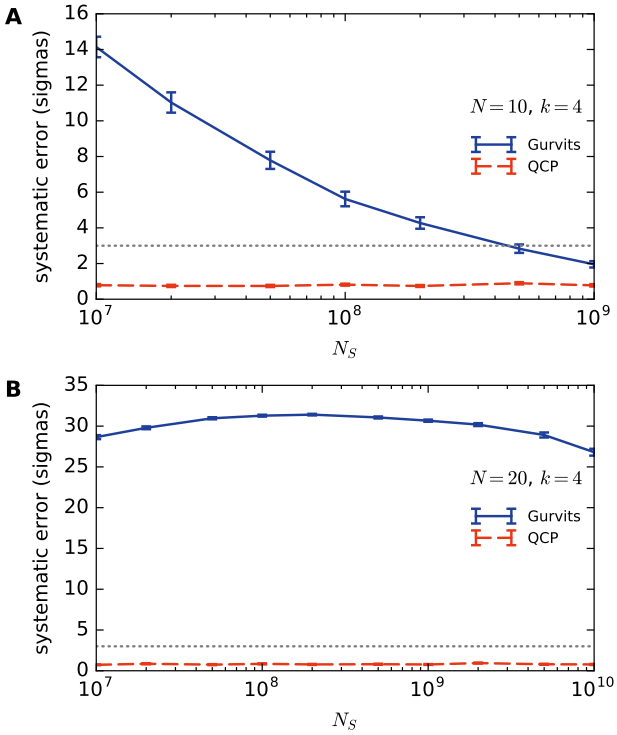

FIG. 2. **Systematic errors for the modulus-squared of the permanent**. **(A-B)** Systematic error of the Gurvits method and the QCP representation with $d \to \infty$, expressed in the units of the sampling error, as a function of the total number of samples $N_s = N_e N_{se}$. For each point we have used $N_m = 100$ random unitary matrices and $N_e = 1000$ ensembles of $N_{se} = 10^4 \ldots 10^6$ (10^7 for the panel **B**) sub-ensemble sizes. Solid blue lines correspond to the Gurvits method, dashed red lines corresponds to the QCP method and dotted grey lines correspond to the confidence interval (3 sigmas). Here we consider $k = 4$, (**A**) $N = 10$ and (**B**) $N = 20$.

Our method gives an unbiased estimate of the permanent squared, which is the relevant quantity in boson sampling experiments. Even more importantly, our methods can also be used in the practical case that the input is not a pure binary number state, or where the output measurements are not n-th order correlations.

From the mathematical viewpoint, these numerical methods unify the complex P-representation approach in quantum optics with the computational problem of efficient approximations of permanents and the permanent squared, which is the relevant calculation for boson sampling.

In order to show that these novel complex P-representation methods are also more efficient than other phase space methods, in Fig. 3, we plot the relative sampling errors of the permanent squared averaged over random matrices using the VCP representation with $r = 0.1$, the QCP with $d \to \infty$ and the Q-function [28].

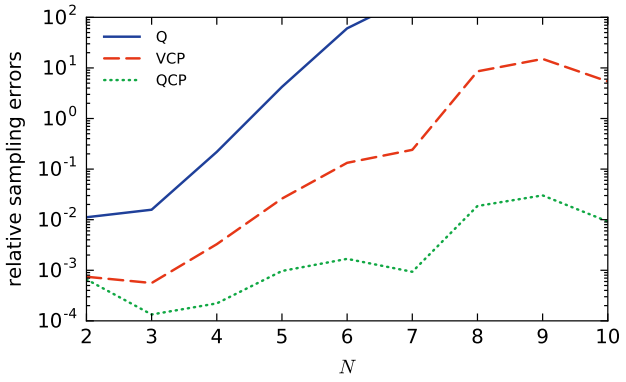

FIG. 3. **Sampling errors for different phase space methods.** Relative sampling errors of the permanent squared (the maximum-order nonzero correlation) as a function of the sub-matrix size N using the Q-function, the QCP with $d \to \infty$ and the VCP method with a parameter $r = 0.1$. Here we used $k = 2$, $N_e = 1000$ ensembles with $N_{se} = 10^5$ sub-ensemble size and $N_M = 10$ random matrices for each point.

Combined correlations verification strategy

As these experiments improve, the problem of verification at large N will become acute [12, 29–34]. One needs a verification protocol that is calculable and measurable at large N, is related to permanents, depends on the unitary matrix, can distinguish different unitaries, and cannot be mimicked by other processes. Since the exact permanents are exponentially hard to compute, an ideal solution is to have an analytic signature that is measurable.

To obtain a large N quantity that is both measurable and calculable, we first consider the *combined* correlation $C_{N|M} \equiv \langle \hat{C}_{N|M} \rangle$, defined as the sum over all different channel combinations σ of length N:

$$\hat{C}_{N|M} = \sum_{\sigma} \prod_{p=1}^{N} \hat{n}_{\sigma_p}. \tag{5}$$

These sums combine an exponentially large number of permanents, each of which is exponentially hard to compute. Despite this, they can be evaluated efficiently using the QCP method using a modification of this technique that utilises a discrete Fourier transform (DFT); see Materials and Methods for details. We note that while Gurvits type binary methods can also be used in combination with the DFT approach, the sampling error is greatly improved with a qudit or continuous sampling.

This provides an exceptionally efficient route to the randomised calculation of *all* the exponentially large number of N-th order correlations, each involving an exponentially large permanent calculation. The advantage of using $C_{N|M}$ is that it allows to calculate a signature that has high count rates even at large N values, and therefore is scalable. This can be measured experimentally with high average count rates, especially in the important large k regime [26, 27].

However, there is a cost involved: this is still not exact for finite resources, due to sampling errors both in the experiment measurements and theoretical estimates. The difficulty is that, for large N, $C_{N|M}$ is almost always given by its unitary average. Although satisfying the other criteria, it doesn't adequately distinguish between the different unitaries due to sampling errors in the measurement process. A fraudulent boson sampler could therefore be constructed to replicate the required statistics, without having to process any information about the unitary.

Channel-deleted verification

An improved strategy is needed in order to discriminate between different unitary matrices, and obtain a unique signature of the required permanent distribution. Therefore, to verify the boson sampler, we now consider a hierarchy of experiments in which one or more channels is deleted from the channel combinations. This leads to channel-deleted, combined correlation $C^{(p)}_{N|M}$, which sum over all N-fold correlations that *don't* include a specific channel p:

$$\hat{C}^{(p)}_{N|M} = \sum_{\sigma, p \notin \sigma} \prod_{j=1}^{N} \hat{n}_{\sigma_j}. \tag{6}$$

In this approach we measure the combined permanents conditioned on channel p having no counts. Similarly, one can have a hierarchy of measurements $C^{(p,q,\ldots)}_{N|M}$, which are conditioned on channels $p, q, \ldots$ have no counts. Eventually this exhaustively enumerates all the possible N-th order correlations, starting with the most readily measured combinations having the highest count rates and the lowest experimental errors.

The advantage of this approach is that the goal of a boson sampling device is to generate samples with a permanent distribution. However, any probability distribution over a finite range has a unique fingerprint [35]: the set of all its moments. The hierarchy of channel-deleted combinations converges to this unique signature, in the limit in which all possible deletions are included.

The specified channel becomes a loss reservoir for the remaining $M-1$ channels. This modified correlation can be readily evaluated using the QCP method, since its expectation value can be calculated probabilistically as:

$$C^{(p)}_{N|M} = \left\langle \frac{1}{M-1} \sum_{j} e^{-ijk\delta} \prod_{k=1}^{M-1} \left[1 + e^{ij\delta} m_k\right] \right\rangle. \tag{7}$$

Here m_k is a scaled boson number. We note, however, that even this technique eventually has too large a sampling error for large enough N values. Eventually, it will

become essential to obtain an analytically calculable signature to verify boson sampling for the large N case, which is important from the complexity theory viewpoint.

For this modified correlation with a single channel-deletion we give a physical argument that allows to conjecture an analytical form at large N. In this case the selected channel p acts as a loss reservoir for the other $M-1$ channels. Therefore, we can write $\alpha^{(\mathrm{out})} = U\alpha^{(\mathrm{in})}$, $\beta^{(\mathrm{out})} = \beta^{(\mathrm{in})}U^\dagger$ so that

$$\left\langle n_j^{(\mathrm{out})} \right\rangle = \left\langle U\alpha^{(in)}\beta^{(in)}U^\dagger \right\rangle_{jj} = \sum_{i=1}^{N} |U_{ij}|^2 = T_j,$$

and $\sum_{j=1}^{M} T_j = N$. Next, we exclude the counts in channel p, since this acts as a loss reservoir. From unitarity, $\sum_{j=1}^{M} T_j = Nt$, and we get:

$$\sum_{j\neq p}^{M} T_j = Nt - T_p = Nt\,(1 - T_p/Nt) = t_p N,$$

where the effective channel loss rate for channel p is $t_p = t\left(1 - \sum_{i=1}^{N} |U_{pi}|^2 / N\right)$.

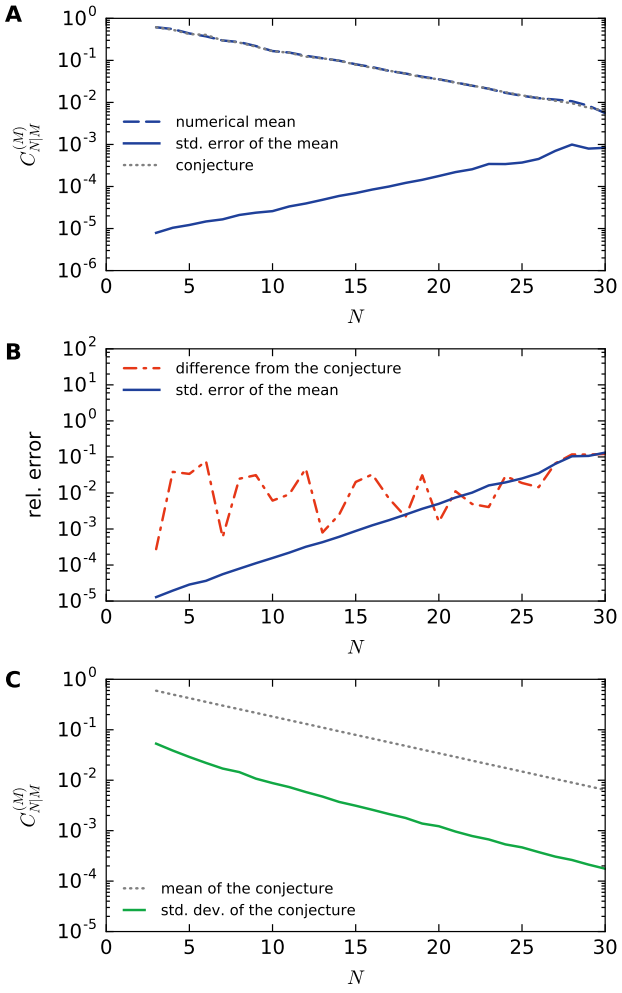

FIG. 4. **Combined correlations**. (**A**) Combined correlations $C^{(M)}_{N|M}$ given in Eq. (7) evaluated using the QCP method. Here we have used $d \to \infty$, $k = 6$, $N_m = 1$ and $N_e = 10^4$ ensembles of 10^6 samples. The dotted grey line is the analytical result given in Eq. (8), where $N = M/k$. The dashed blue line corresponds to our numerical results. The solid blue line is the average estimate of the error in the mean $S = \langle\sqrt{\langle C^2\rangle - \langle C\rangle^2}\rangle_m/\sqrt{N_e}$. (**B**) Modulus of the difference between the numerically obtained combined correlation, and the conjecture $C^{(M),\mathrm{conj}}_{N|M}$ (dash-dotted red line) plotted against the error estimate from the top panel, normalised on $C^{(M),\mathrm{conj}}_{N|M}$. (**C**) Standard deviation of the analytic conjecture for 10^4 random matrices (solid green line) plotted against its mean value, showing that there is a substantial variation between unitaries even for relatively large N.

For large $k \equiv M/N$, the modified correlation — which is the probability of N coincidences in any of the remaining $M-p$ channels — is given by the sum of permanents of the $N \times N$ sub-matrices of the remaining $(M-1)\times(M-1)$ matrix. We conjecture that this is given asymptotically by taking an $M-1$ dimensional unitary average, together with an additional reduction

of t_p^N due to channel loss, in channel p, so that:

$$C_{N|kN}^{(p),\mathrm{conj}} \underset{k\to\infty}{\sim} \frac{t_p^N (kN-1)!(kN-2)!}{((k-1)N-1)!((k+1)N-2)!}, \quad (8)$$

We have tested this result using quantum software methods, that allow us to sum over exponentially large numbers — up to 10^{34} — of large permanents in parallel. In Fig. 4 we show numerical results for the channel combinations $C_{N|M}^{(p)}$ using the QCP method as well as for the conjecture $C_{N|M}^{(p),\mathrm{conj}}$. This remarkable task of summing over exponentially large numbers of quantities, each of which is exponentially difficult to compute, would take some trillions of lifetimes of the universe using previously known exact or approximate methods.

Our numerical simulations show that the conjecture is asymptotically valid at large k for a single unitary, apart from highly anomalous matrices with zero asymptotic measure, such as the unit matrix. It is in excellent agreement with our quantum simulations of a random unitary to a relative accuracy of order 1% for $k = 6$. This signature is purely analytic, and calculable for large sizes. It is experimentally accessible, and can distinguish unitaries. It can be generalised to arbitrary channel deletions, giving an increasingly unique fingerprint of the permanent distribution.

This is more powerful than methods that distinguish a boson sampler from one with a uniform distribution [12, 30, 31] or distinguish particles based on their statistics [29]. It is more scalable than methods requiring knowledge of the permanent [1], including coarse-graining [30], because permanents cannot be computed exactly for large matrices. It is not restricted to Fourier-transform unitaries [32]. It is applicable, with modifications, to recently proposed scatter-shot experiments [33, 36]. Efficient nanowire detectors with $t = 0.9$ (using an optimistic estimate) and 1 GHz measurement rates [37] could allow one to reach count rates as high as one per second even for $N = 100$ and $k = 10.$, so this appears potentially viable in terms of count rates.

Simulating QuFTI

Finally, to test our methods for quantum enhanced technology we consider the following quantum metrology experiment: the quantum Fourier transform interferometer (QuFTI) [3]. This is a novel multi-mode interferometer that measures gradients of a phase-shift and corresponds to consider the case where the input state is one photon in each input mode. Here the permanent of the full matrix is evaluated, not the permanents of submatrices, so that $M = N$. Motes *et al* [3] conjecture that the count rate, $Q^{(\mathrm{conj})}$, showing an M-fold improvement in fringe visibility is:

$$Q^{(\mathrm{conj})} = \left|\mathrm{perm}U^{(M)}\right|^2 = \frac{1}{M^{2M-2}} \prod_{j=1}^{M-1} \left[2j(M-j)\cos(M\phi) + M^2 - 2jM + 2j^2\right]. \quad (9)$$

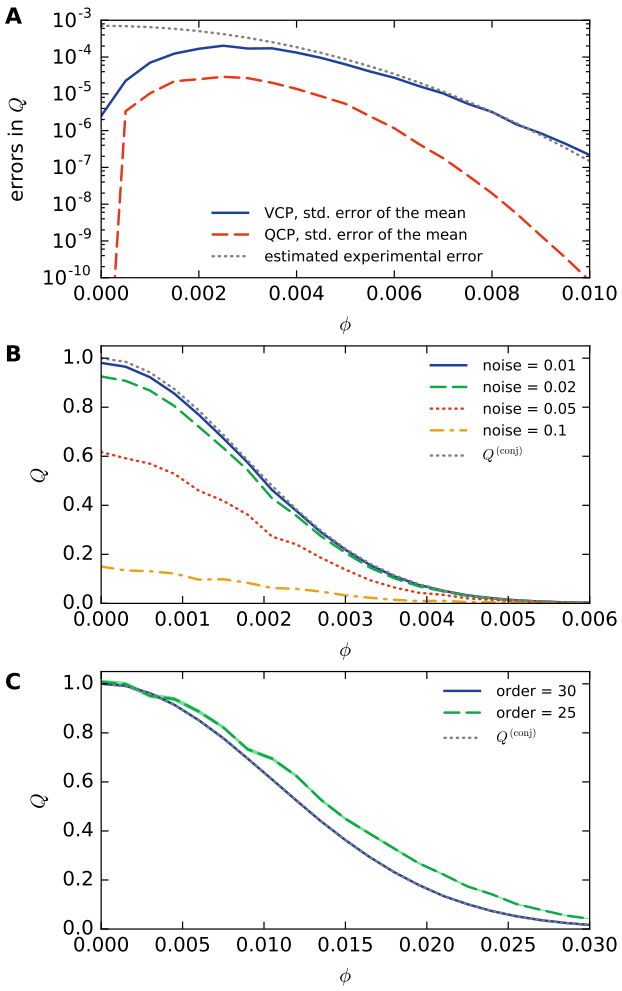

FIG. 5. **Quantum Fourier transform interferometer (QuFTI)**: (**A**) Estimated error in the mean of modulus square of the permanent denoted by Q, $\sqrt{\langle Q^2\rangle - \langle Q\rangle^2}/\sqrt{N_e}$, using two different methods of the quantum software. The solid blue line corresponds to the VCP with $r = 0.1$, while the red dashed line corresponds to the QCP with $d = 2$. The dashed grey line is the estimated experimental error $\sqrt{Q^{(\mathrm{conj})}/(N_e N_s)}$ ($N_e = 200$, $N_s = 10^4$). Here $Q^{(\mathrm{conj})}$ is given in Eq. (9). (**B**) Value of the interferometry correlation Q, for different amounts of noise included in the measured angle. The values were obtained using QCP with $d = 2$. Here we have considered $M = 100$, $N_e = 200$ ensembles, $N_s = 10^4$, the results are averaged over 20 matrices. (**C**) Correlations of different order for $M = 30$ obtained using the VCP method with 200 ensembles of 10^6 samples. The VCP parameter was set to $r = 0.8$ to minimise the sampling error (see Fig. 6b).

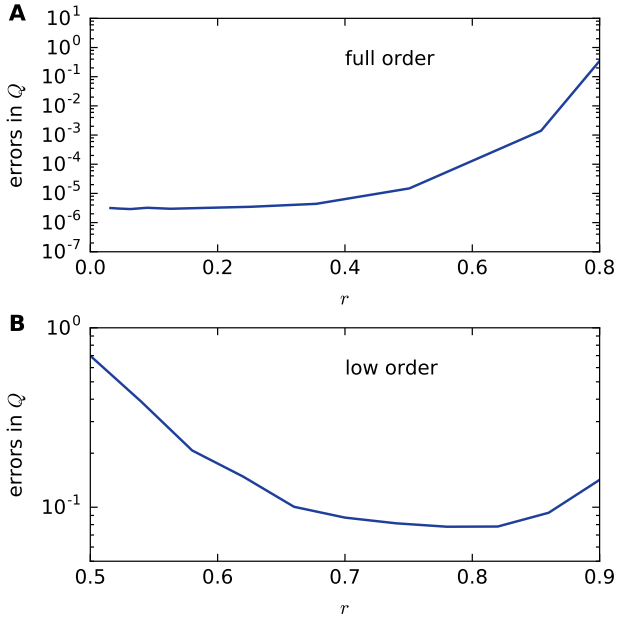

FIG. 6. **Scaling of the error for the VCP parameter:** (**A**) Dependence of the estimated error in the mean with respect to the parameter r for the VCP method. Here we have used $M = 100$ and $\phi = 0.007$, $N_e = 200$, $N_s = 10^5$. (**B**) Dependence of the error for the 25th order correlation as a function of r for $M = 30$ and $\phi = 0.01$, obtained with the VCP method with 200 ensembles of 10^4 samples.

Fig. 5 shows results obtained with our quantum software for the QuFTI. We confirm the interferometer conjecture with an unprecedented permanent size of 100×100. Exact methods would take trillions of years to compute this. In Fig. 5A we evaluate the error in the modulus square of the permanent using both the VCP and QCP representation with $d = 2$.

In our simulations we include the effect of decoherence, using an independent phase noise term ξ at each site with a normal distribution, having zero mean and variance σ^2. This is vitally important, since other Heisenberg-limited interferometry proposals are fragile against such decoherence effects. Each mode has a noise term given by $\phi_j = j\phi + \xi_j$, $j = 1, \ldots, M$. This effect of the noise is shown in Fig. 5B for different values of the variance. We also evaluate different order correlations; since experimental inefficiencies may lead to greater count rates for lower order correlations. This is shown Fig. 5C, for $M = 30$, where we show three different order correlation functions, $Q = \langle \hat{n}_1 \ldots \hat{n}_M \rangle$. Figure 6 gives the sampling error as a function of r, showing that the optimal contour radius depends on the order of the correlations.

DISCUSSION

In summary, we obtain novel "quantum software" methods based on complex phase-space distributions. These can be used for the simulation of linear photonic networks. Our results are valid for arbitrary inputs, losses and outputs. We calculate how the sampling error scales with the number of photons. We propose a combined channel grouping protocol that allows one to validate large-scale boson sampling experiments. Our results also demonstrate the exceptional robustness of N-photon quantum Fourier interferometry against decoherence.

These results are limited by sampling errors, even with the highly efficient methods we use. Since the analytic test we propose is a conjecture, its limitations need to be investigated by further studies using random matrix theory tools.

Acknowledgments: This research was supported in part by the National Science Foundation under Grant No. NSF PHY-1125915, and by the Australian Research Council. We also wish to acknowledge discussions and correspondence with Scott Aaronson, Yan Fyodorov, Peter Rohde, Fabio Sciarrino, Christine Silberhorn and Ian Walmsley.

MATERIALS AND METHODS

Circular von Mises

In order to sample the initial state according to the VCP distribution(2), we split it into a probability distribution and a complex weight that is used for calculating moments.

The sampled probability distribution is:

$$P(\boldsymbol{\theta}, \boldsymbol{\phi}) = \prod_k \frac{1}{2\pi I_0(r^2)} \exp\left(r^2 \cos\theta_k\right),$$

where $I_0(x)$ is the modified Bessel function of the first kind of order 0, $\theta_k \in [-\pi, \pi)$, and $\phi_k \in [-\pi, \pi)$. This corresponds to sampling the variables separately as:

$$\theta_k = \mathcal{U}(-\pi, \pi),$$

$$\phi_k = \mathcal{VM}\left(0, r^2\right),$$

where $\mathcal{U}$ is the uniform distribution, and $\mathcal{VM}$ is the circular von Mises distribution [22]. For each sample we calculate the phase-space variables as $\alpha_k = r\exp(i(\phi_k + \theta_k/2))$, $\beta_k = r\exp(i(\phi_k - \theta_k/2))$. For factors where $n_k = 0$ we use the optimal distribution $P(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_k \delta(\alpha_k)\delta(\beta_k)$ instead, that is, take $\alpha_k = \beta_k = 0$.

The corresponding complex weights are

$$\Omega = \prod_k \begin{cases} \frac{(n_k!)^2 I_0(r^2)}{r^{2n_k}} \exp\left(i\left(r^2 \sin\theta_k - n_k\theta_k\right)\right), & n_k > 0, \\ 1, & n_k = 0. \end{cases}$$

These are used to calculate a moment $f(\boldsymbol{\alpha},\boldsymbol{\beta})$ in conjunction with the drawn samples of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as

$$\langle f(\boldsymbol{\alpha},\boldsymbol{\beta})\rangle = \frac{1}{N_s}\sum_{j=1}^{N_s} \Omega\left(\boldsymbol{\alpha}^{(j)},\boldsymbol{\beta}^{(j)}\right) f\left(\boldsymbol{\alpha}^{(j)},\boldsymbol{\beta}^{(j)}\right),$$

where N_s denotes the number of samples, and $\boldsymbol{\alpha}^{(j)}$ and $\boldsymbol{\beta}^{(j)}$ is the sampled pair of random phase-space coordinates.

Qudit sampling

The complex qudit P-distribution (QCP) can be shown to exist using discrete Fourier transforms. The distribution in this case corresponds to a discrete choice of phases ϕ_k, allowing one to obtain more efficient sampling in some cases, especially in Fourier transform interferometry.

The result is similar to a path integral, except that the distribution is now a discrete sum over delta-functions in phase. The complex qudit P-function P_Q is:

$$P_Q(\mathbf{q},\tilde{\mathbf{q}}) = \sum_{\boldsymbol{n},\boldsymbol{m}} \langle \boldsymbol{m}|\,\widehat{\rho}\,|\boldsymbol{n}\rangle \prod_j r^{-(n_j+m_j)} e^{i\phi(n_j\tilde{q}_j - m_j q_j)}\sqrt{n_j! m_j!}.$$

This can be verified as a solution, by inserting this distribution into the expansion of the density matrix, and noting that, from the properties of the discrete Fourier transform,

$$\frac{1}{d^M}\sum_{\mathbf{q}} e^{i\boldsymbol{q}\cdot(\boldsymbol{n}-\boldsymbol{m})\phi} = \delta_{\boldsymbol{n}-\boldsymbol{m}}.$$

In the one mode, one boson case, then apart from a radial factor, the P-function is simply the product of two complex variables:

$$P_Q(q,\tilde{q}) = \frac{1}{r^2}\tilde{z}^{\tilde{q}}\left(z^q\right)^* .$$

The advantage of the small radius limit is that the known identities for the generalised P-representation are all valid, provided the $r \to 0$ limit is taken after the calculation. Thus, we can use the standard result that after transmission through a linear optical system with phase-shifts, beam-splitters and losses, the output coherent amplitudes are multiplied by the relevant linear transmission matrix. In calculating N-th order correlations of an N-photon input, all the factors proportional to the radius r simply cancel.

The photon number phase-space variable, for an input of single bosons into the first N modes, is

$$n_k^{(\mathrm{out})}(q,\tilde{q}) = r^2\left(\sum_{j\in S} T_{kj} z_j^q\right)\left(\sum_{j'\in S} T_{kj'} z_{j'}^{\tilde{q}}\right)^* ,$$

As a result, after including the complex P-function weights, an N-th order output correlation is

$$\left\langle \prod_{k\in S}\hat{n}_k\right\rangle_Q = \left|\frac{1}{d^M}\sum_{\mathbf{q}}\left\{\prod_{i\in S} z_i^{q*}\prod_{k\in S'}\left(\sum_{j\in S} T_{kj} z_j^q\right)\right\}\right|^2 .$$
$$= \left|\mathrm{perm}\left[T\left(S',S\right)\right]\right|^2$$

As expected [5], this is the square of the permanent of the sub-matrix of T with rows in S' and columns in S, which we call $T\left(S',S\right)$. After summation on the $\mathbf{q}$ indices, the only terms that survive involve products of distinct permutations of the matrix indices, which is the permanent. There are exponentially many terms involved at large N. For reasons of efficiency, this is evaluated computationally by taking randomly chosen integers $(\mathbf{q},\tilde{\mathbf{q}})$, and averaging over many samples of these random phases.

Evaluation of sampling errors for the Gurvits and QCP method

The data processing for Fig. 2 is performed as follows. Each data point in the plot corresponds to N_s total samples, which is split in $N_e = 10^3$ ensembles of $N_{se} = N_s/N_e$ samples. For each data point we obtain a $N_m \times N_e'$ matrix $P_j^{(m)}$ of permanent squared values for each of N_m random matrices and each ensemble. The effective number of ensembles N_e' is equal to N_e for the Gurvits method, and $N_e/2$ for the QCP method (since two independent ensembles are used to produce a single ensemble value of the permanent squared). We also calculate the exact permanent squared values $\tilde{P}^{(m)}$ for each random matrix.

Per-matrix systematic errors and sampling errors are calculated as:

$$D^{(m)} = \left|\frac{1}{N_e'}\sum_{j=1}^{N_e'} P_j^{(m)} - \tilde{P}^{(m)}\right| \equiv \left|\langle P^{(m)}\rangle_e - \tilde{P}^{(m)}\right|,$$

$$E^{(m)} = \frac{\sqrt{\langle\left(P^{(m)}\right)^2\rangle_e - \langle P^{(m)}\rangle_e^2}}{\sqrt{N_e'}}.$$

We then plot the systematic errors D measured in the respective sampling errors E, that is, we calculate the values

$$S^{(m)} = D^{(m)}/E^{(m)}$$

and plot the mean $\langle S\rangle_M$ and the estimated error of the mean $\sqrt{\langle S^2\rangle_M - \langle S\rangle_M^2}/\sqrt{N_M}$ as error-bars.

Combined correlations

We use the QCP method in order to evaluate the sums of combined correlations $C_{N|M} \equiv \langle \hat{C}_{N|M}\rangle$ of the form given by

$$\hat{C}_{N|M} = \sum_{\sigma} \prod_{j=1}^{N} \hat{n}_{\sigma_j}.$$

These sums combine an exponentially large number of permanents, each of which is exponentially hard to compute. We first consider the following type of normally ordered correlation polynomial, calculated for $j = 0, \ldots M-1$:

$$\hat{D}_j = \prod_{k=1}^{M} \left[1 + e^{ij\delta}\hat{n}_k\right],$$

with $\delta \equiv 2\pi/M$. There are M distinct correlation polynomials for $j = 0, \ldots M-1$, each including all possible combinations of $\hat{n}_k$. The inverse Fourier transform is given by:

$$\hat{F}_k = \frac{1}{M}\sum_j e^{-ijk\delta}\hat{D}_j.$$

When $k = N$, the phase factors for N-th order correlations cancel, and we get the following result, that includes all the output correlations in all sets S' of length N:

$$\hat{F}_N = \sum_{S'} \prod_{k\in S'} \hat{n}_k \equiv C_{N|M}\left(\hat{\mathbf{n}}\right).$$

We now wish to relate these combined correlations to the evaluation of the modulus squared of the permanent of all the unitary sub-matrices $\mathbf{M}$. Hence we consider the case where $\boldsymbol{\alpha}^{(\mathrm{out})} = \boldsymbol{T}\boldsymbol{\alpha}^{(\mathrm{in})}$, and $\boldsymbol{T}$ is a unitary matrix. As described above, we use random variables z_i with unit modulus, and the Einstein summation convention. A stochastic mean over the random phase variables is denoted $\langle\rangle$.

For each sub-matrix $\mathbf{M}$, we know that:

$$\mathrm{perm}\left(\mathbf{M}\right) = \left\langle \prod_i z_i^* \prod_k \left(M_{kj} z_j\right)\right\rangle.$$

Using the QCP method we can evaluate the modulus square of the permanent as in Eq. (4), obtaining:

$$|\mathrm{perm}(\mathbf{M})|^2 = \left\langle \prod_i z_i^* \prod_k \left(M_{kj} z_j\right) \prod_i z_i' \prod_k \left(M_{kj} z_j'\right)^*\right\rangle$$

We note that for measurements in the selected channels of the sub-matrix, $\alpha_k = rM_{kj}z_j$, $\beta_k^* = rM_{kj}z_j'$. Introducing a scaled boson number $m_k = \alpha_k\beta_k/r^2$, and a combined unit random $y_i = z_i^* z_i'$, gives the following expression:

$$|\mathrm{perm}(\mathbf{M})|^2 = \left\langle \prod_j y_j \prod_k m_k\right\rangle.$$

For the present problem, in terms of permanents of a sub-matrix $\tilde{T}$ of size $S' \times S$ which will be denoted by $\tilde{T}\left(S', S\right)$, we obtain:

$$\sum_{S'} \left|\mathrm{perm}(\tilde{\mathbf{T}}\left(S, S'\right))\right|^2 = \sum_{S'} \left\langle \prod_{j\in S} y_j \prod_{k\in S'} m_k\right\rangle.$$

Next, using the Fourier expansion result given above, we get:

$$\sum_{S'} \left|\mathrm{perm}(\tilde{\mathbf{T}}\left(S, S'\right))\right|^2 = \left\langle C_{N|M}\left(\mathbf{m}\right) \prod_{j\in S} y_j\right\rangle,$$

with:

$$C_{N|M}\left(\mathbf{m}\right) = \frac{1}{M}\sum_j e^{-ijN\delta} \prod_{k=1}^{M}\left[1 + e^{ij\delta}m_k\right].$$

We note that the above expression allow us to perform the randomised calculation of *all* the exponentially large numbers of N-th order correlations. This would not be feasible with any previous algorithm that does not use the innovative Fourier transform strategy.

[1] S. Aaronson and A. Arkhipov, in *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing* (ACM Press, 2011) pp. 333–342.
[2] S. Aaronson and A. Arkhipov, Theory of Computing **9**, 143 (2013).
[3] K. R. Motes *et al.*, Phys. Rev. Lett. **114**, 170802 (2015).
[4] J. Carolan *et al.*, Science **349**, 711 (2015).
[5] S. Aaronson, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **467**, 3393 (2011).
[6] L. Valiant, Theoretical Computer Science **8**, 189 (1979).
[7] M. A. Broome *et al.*, Science **339**, 794 (2013).
[8] A. Crespi *et al.*, Nat. Photon. **7**, 545 (2013).
[9] M. Tillmann *et al.*, Nat. Photon. **7**, 540 (2013).
[10] J. B. Spring *et al.*, Science **339**, 798 (2013).
[11] A. Crespi *et al.*, Nat. Commun. **7**, 10469 (2016).
[12] N. Spagnolo *et al.*, Nat. Photon. **8**, 615 (2014).
[13] J. C. Loredo *et al.*, https://arxiv.org/abs/1603.00054 (2016).
[14] S. Scheel, http://arxiv.org/abs/quant-ph/0406127 (2004).
[15] S. Scheel, in *Quantum Information Processing*, edited by T. Beth and G. Leuchs (Wiley-VCH, Weinheim, 2005) Chap. 28, pp. 382–392.

[16] D. G. Glynn, European Journal of Combinatorics **31**, 1887 (2010).

[17] J. Wu, Y. Liu, B. Zhang, X. Jin, Y. Wang, H. Wang, and X. Yang, ArXiv e-prints (2016), arXiv:1606.05836.

[18] P. D. Drummond and C. W. Gardiner, J. Phys. A **13**, 2353 (1980).

[19] Y. He *et al.*, http://arxiv.org/abs/1603.04127 (2016).

[20] J. P. Olson *et al.*, http://arxiv.org/abs/1610.07128 (2016).

[21] P. D. Drummond and M. Hillery, *The Quantum Theory of Nonlinear Optics* (Cambridge University Press, 2014).

[22] K. V. Mardia and P. E. Jupp, *Directional statistics*, Wiley series in probability and statistics (John Wiley and Sons, 2000).

[23] Y. V. Fyodorov, International Mathematics Research Notices **2006**, 61570 (2006).

[24] P. D. Drummond and M. D. Reid, arXiv:1610.01261 (2016).

[25] L. Gurvits and A. Samorodnitsky, Discrete & Computational Geometry **27**, 531 (2002).

[26] A. Arkhipov and G. Kuperberg, Geometry & Topology Monographs **18**, 1 (2012).

[27] P. D. Drummond, B. Opanchuk, L. Rosales-Zárate, M. D. Reid, and P. J. Forrester, Phys. Rev. A **94**, 042339 (2016).

[28] K. Husimi, Proc. Phys. Math. Soc. Jpn **22**, 264 (1940).

[29] M. Walschaers *et al.*, New Journal of Physics **18**, 032001 (2016).

[30] J. Carolan *et al.*, Nat. Photon. **8**, 621 (2014).

[31] S. Aaronson and A. Arkhipov, Quantum Info. Comput. **14**, 1383 (2014).

[32] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Mølmer, Phys. Rev. Lett. **113**, 020502 (2014).

[33] M. Bentivegna *et al.*, Science Advances **1**, e1400255 (2015).

[34] L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert, Nat. Commun. **6**, 8498 (2015).

[35] P. A. P. Moran, *An introduction to Probability Theory* (Oxford, Clarendon Press, 1968).

[36] A. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. OÕBrien, and T. Ralph, Phys. Rev. Lett. **113**, 100502 (2014).

[37] H. Takesue, S. D. Dyer, M. J. Stevens, V. Verma, R. P. Mirin, and S. W. Nam, Optica **2**, 832 (2015).